**IN THE UNITED STATES**

**PATENT AND TRADEMARK OFFICE**

| | |
|---|---|
| APPLICANTS: | Adam T. Zemla |
| APPLICATION NO.: | 10/782,061 |
| FILING DATE: | February 18, 2004 |
| TITLE: | Local-Global Alignment for Finding 3D Similarities in Protein Structures |
| EXAMINER: | Michael L. Borin |
| GROUP ART UNIT: | 1631 |
| ATTY. DKT. NO.: | IL-11160 |
| F&W REF: | 26303-13766US |

**DECLARATION OF CAROL ECALE ZHOU, PH.D. UNDER 37 C.F.R. § 1.132**

Sir:

I, Carol Ecale Zhou, Ph.D., hereby declare as follows:

1. I am a Scientist at Lawrence Livermore National Laboratory in Livermore, California in the Energy, Environment, and Biology Division L-174. I received my bachelor's degree in chemistry from Purdue University. I completed my doctoral training in biological science and entomology at the University of Missouri in Columbia, Missouri. In addition, I received a bachelor's degree in computer science from California State University in Sacramento, California. A true and correct copy of my *Curriculum Vitae* is attached to this declaration as Exhibit A. If called as a witness I could competently testify to the facts and opinions expressed in this declaration.

2. My primary research at Lawrence Livermore National Laboratory relates to pathogen bio-informatics. I have served as a principal investigator and an administrative group leader in the Pathogen Bio-informatics Group at Lawrence Livermore National Laboratories. Pathogen bio-informatics is an area of science directed to the analysis of biological sequence and protein structure data from pathogens, e.g., organisms considered to be agents of biological warfare and disease.

3. For the last six years I have developed protein structure bio-informatics algorithms and applied these algorithms to pathogen data for bio-terrorism defense research. I am the author of several peer-reviewed publications, listed in my *Curriculum Vitae*.

4. I have reviewed pending U.S. patent application 10/782,061 ("the '061 application") and am familiar with the methods described and claimed therein for "generating a local-global alignment score which indicates a global and a local similarity between a first protein structure and a second protein structure."

5. The local-global alignment score described and claimed in the '061 patent is a protein structure homology score as it "indicates a global and a local similarity between a first and a second protein structure." Protein structure homology scores are well-known to one of skill in the art and are used to characterize similarity between two protein structures in order to identify homologous proteins. Homologous proteins are proteins that are related due to the processes of speciation, genetic engineering and/or duplication. Due to structural conservation, two homologous protein structures retain similar structures in order to conserve the functional properties of their respective protein structures. Therefore, protein structure homology scores are used to characterize both the structural and functional similarities of proteins. The structures of homologous proteins are less likely to diverge during evolution compared to the protein or

DNA sequences of homologous proteins. Therefore, protein structure homology scores are regarded by those skilled in the art as a "gold standard" in identifying homologous proteins.

6. As defined in 2107.01 of the Manual for Patent Examining Procedure, "to satisfy a substantial utility an asserted use must show that the claimed invention has a significant and presently available benefit to the public." *Fisher*, 421 F.3d at 1371, 76 USPQ2d at 1230. It is further defined that the substantial utility cannot require carrying out further research to identify or reasonably confirm a "real world" context of use.

7. Protein structure homology scores are used to assign various functional characteristics to unknown proteins. As described in more detail below, this information is used in a number of settings, including public health and drug discovery, to make decisions related to those fields. Therefore it is my opinion that methods described and claimed in the '061 patent satisfy the requirements for substantial utility as they have a utility that is significant and presently available, as is readily apparent to those skilled in the art. It is also my opinion that the methods described and claimed in the '061 patent do not require further research in order to identify or confirm this real-world utility. If a protein structure homology score indicates a level of similarity between two protein structures that is much larger than what could be expected by chance, then the proteins are identified as homologs and functional characteristics of one protein structure can be assigned to another without further experimental validation.

8. One example of the utility of protein structure homology scores is provided by the field of pathogen detection and identification. For example, protein structure homology scores have many practical applications in bioterrorism defense research. In bioterrorism defense research, functional characteristics such as virulence, drug resistance and toxicity of a known protein are assigned to a new or novel protein structure based on protein structure homology scores

indicating a high level of structural similarity. Protein structure homology scores are used to identify unknown pathogens based on the similarity of protein structures of a known protein and a protein isolated from the unknown pathogen. Tactical military and public health decisions are made based on the identification of these proteins. For example, if a protein structure for a protein thought to be produced as an agent of bio-terrorism has a protein structure homology score which indicates that it is a genetically modified version of an existing pathogen such as Ricin, national security decisions can be made based on the protein homology score without further experimental validation.

9. Some examples of these applications are outlined in detail in Slezak et al.,"Comparative genomics tools applied to bioterrorism defense," *Briefings in Bioinformatics*, 2003, 4(2), 133-149 ("the Slezak reference"). A copy of the Slezak reference is submitted with the concurrently submitted IDS. On page 142, paragraph 2 of the Slezak reference the authors describe the use of protein homology in computational modeling of proteins in order to assess signatures used in fieldable detection assays: "*Our pipeline has generated numerous signatures for multiple bacterial and viral pathogens that have proven to work well in extensive field use. Many of these signatures were 'anonymous' when they were created, meaning that little, if anything, was known about the region upon which they landed. To find out more, a homology-based computational protein structure modelling system is being developed.*"

10. Protein structure homology scores also have practical utility in drug design. In drug design, the identification of protein structure homologs from two organisms such as, e.g., a host organism and a pathogen, is used to identify proteins likely to have cross-reactivity with therapeutics. The identification of protein structure homologs between a host organism and a pathogen is used to target therapeutic development to protein structures in the pathogen with no

cross-reactivity with the host organism. Conversely, the identification of homologous protein structures within a large group of pathogens using protein structure homology scores is used to identify targets for broad spectrum drug development.

11. The validity of the use of protein structure homology scores in drug design is fully described in the art, including the following publications: Takeda-Shitaka et al., "Protein Structure Prediction in Structure based Drug Design", Current Medicinal Chemistry, 2004, 11(5), 551-558 ("the Takeda-Shitaka reference"), Legnauer et al., "Protein Structure Prediction Methods for Drug Design", Briefings in Bioinformatics, 2000, 1(3) 275-288 and Kopp et al., "The SWISS-MODEL Repository of annotated three-dimensional protein structure homology models", 2004, 1(32), D230–D234 ("the Kopp reference"). For example, page 277, paragraph 3 of Legnauer discusses the use of protein homology to determine functional characteristics of proteins for drug development: "*If the homology is above, say, 40 per cent and functionally important motifs are conserved then we can hypothesize that the query sequence has a function that is quite similar to that of the model sequence. As the level of similarity decreases, the conclusions on function that can be drawn from sequence similarity become less and less reliable.*" Copies of these references are submitted with the concurrently submitted IDS.

12. In conclusion, based on my research experience and my knowledge of the literature available at the time the '061 application was filed, it is my considered opinion that a person having ordinary skill in the art of bio-informatics, working at the time the invention was made, would have readily understood the utility of the invention described and claimed in the '061 application.

13. I further declare that all statements made herein of my own knowledge are true and all statements made on information and belief are believed to be true; and further that these

statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Date: 19 March 2008 By: Carol L. Eisele Zhou

# APPENDIX A: CAROL ZHOU CURRICULM VITAE

**Carol L. Ecale Zhou**
**2380 Corte De La Jara**
**Pleasanton, CA 94566**
**925-846-0907**
**carol.zhou@comcast.net**
**zhou4@llnl.gov**

**Objective: a management position in bioinformatics research & development.** Professional Interests: bioinformatics, software system design; medical countermeasures; pathogen virulence, genomics.

**Professional Experience**

**Computer Scientist**, Lawrence Livermore National Laboratory, Livermore, CA, 9/2000-present. I have been involved in bioinformatics research and development projects as PI, co-PI, technical lead, or team member in the Pathogen Bio-informatics Group in Livermore and at the Joint Genome Institute in Walnut Creek. I have also served as administrative Group Leader for 3-1/2 years, supervising up to 15 software developers. I have demonstrated success in acquiring internal and external funding for research and development efforts, in establishing collaborations, in motivating team members to perform at high levels of productivity, in delivering quality solutions to my customers, in software development in C/C++ and Perl, in relational database schema design, in publishing peer-reviewed papers, and in working with legal council in preparing patent applications.

Projects within the Pathogen Bio-informatics Group at LLNL (3/2002-present):

"Structure-driven Sequence Patterns for Remote Homology Detection" (PI/co-PI; 3-yr project) – Development of novel algorithms for remote protein sequence homology detection. To date our team of 6 has devised a structure-based domain fusion algorithm for detection of protein-protein complexes, and we are working on deriving sequence patterns representative of structurally conserved regions within protein families. I contributed to algorithm development and devised a method for annotation-based validation of templates identified using our domain-fusion algorithm.

"Human Assays: Protein Pipeline" (PI for informatics) – This project was supported by the Dept. of Homeland Security (DHS), and involved the coordinated efforts of scientists at 5 national labs. My role was to collaborate with computational and experimental chemists and other scientists as needed in devising bio-informatics solutions to assist efforts in developing protein-based reagents for identification of Category A pathogens and protein toxins. I devised and implemented sequence- and structure-based methods for identifying protein signatures (conserved/unique regions in proteins). I also designed a high-throughput, whole-proteome sequence and structure analysis system for down-selecting proteins according to assay requirements, which incorporated my codes for identifying protein signatures. I led a team of 4 software developers in building the system over a period of 2-3 years. Intellectual property from this work culminated in several patent applications (see Inventions list, page 3).

"cSpan Algorithm Implementation and Web Interface" (PI) – I led a team of 5 developers in a project involving the development of a web interface to our cSpan algorithm, for identifying functional residues in proteins. cSpan combines structural conservation and structure-alignment-based sequence conservation to identify conserved residues among structurally related proteins.

"Structure Alignment and Classification of Proteins" (team member) – Development of a novel method for automatic structural alignment and clustering of protein structures. Method extends Local-global Alignment (LGA) protein structure analysis software.

"Protein Complex Modeling Infrastructure Evaluation" (PI) – Literature survey of computational and experimental methods for predicting protein-protein complexes. I worked with a specialist in protein structure modeling in a literature survey and evaluation of state-of-the-art methods for protein-complex detection and modeling.

"DNA Signatures Annotation" (Technical Lead) – Designed and implemented a database and software system for DNA signature annotation in support of DHS-funded pathogen detection assay design. I used the system for

annotation of TaqMan signatures for Category A-C pathogens and wrote reports.

Projects at the Joint Genome Institute (9/2000-3/2002): I provided computational support for post-sequencing projects, including the comparison of human chromosome 19 to mouse, and parsing/display of Oak Ridge National Lab bacterial genome annotation data. I wrote web pages for display of the human chromosome 19 gene catalog and microbial annotation data.

Committee Work at LLNL:

- Institutional Laboratory-directed R&D Proposal Evaluation Committee member – FY2007-09
- Computing Applications and Research (CAR) Department Science and Technology Computational Biology Implementation Plan committee chair – August 2007 to date
- Computation Directorate Research Support Committee member – June 2006 to date.
- CAR Department Strategic Planning Group 1.1 member – 2005-2006

**Software Engineer**, Hewlett-Packard Company, Roseville, CA, 1999-2000. C/C++ back-end development of a windows-based application for testing of gigabit routers. The test configuration consisted of three distinct interacting programs running on any number of participating nodes in a local area network; multi-threaded parallel coding; code capabilities included capture of transmission rates and reliability measures of packet transmission; this software system was implemented by a team consisting of myself and two other programmers.

**Postdoctoral Research Associate in Molecular Virology**, Dept. of Entomology, University of California-Davis, 1995-98. Molecular biological and cytological studies of baculoviruses and closteroviruses. I used molecular biology and microscopy techniques to study *Bombyx mori* nucleopolyhedrovirus. Technologies: molecular biological techniques, transmission electron and light microscopy, immunogold labeling, insect cell tissue culture, bio-informatics software for sequence analysis.

**Free-lance Editor**, 1996-97. Writing assistance and substantive editing of manuscripts describing basic research in molecular and cellular biology for Japanese researchers publishing in English language journals.

**Adjunct Professor of Biology**, William Woods University, Fulton, MO, 1994-95. I lectured in general biology and taught an introductory laboratory course, in which I supervised an undergraduate lab aid.

**Graduate Student Researcher, Postdoctoral Research Associate**, Dept. of Entomology, University of Missouri-Columbia, 1988-95. Cytological studies of insect feeding damage and repair in plant vascular tissues. Technologies: light, fluorescence, transmission, and scanning electron microscopy, tissue micro-dissection, plant propagation. I devised a method for applying measurable quantities of insect feeding damage to plant tissues, thereby enabling controlled experimentation and evaluation of plant tissue damage and repair following insect feeding damage.

**AFS International Exchange Teacher**, Novgorod, Russia, 1986. I was assigned by the USSR Ministry of Culture to teach English to 4$^{th}$ year students at the Antonova Pedagogical Institute. In addition to covering State-mandated material, I lectured in American culture and literature. Achieved fluency in spoken Russian language.

**Molecular Biologist**, Sandoz Crop Protection (Zoecon), Palo Alto, CA, 1985-88. Genetic engineering of bio-rational viral insecticides; molecular cloning of foreign gene sequences (e.g., scorpion toxins, insect peptide hormones) into *Autographa californica* nucleopolyhedrovirus. Molecular biology techniques: gene cloning; DNA sequencing; mRNA library construction and screening; Southern/Northern/Western blotting, 2-D gels; insect cell tissue culture, bio-informatics software for sequence analysis.

**College Instructor**, Depts. of Mathematics and Foreign Languages, Purdue University, West Lafayette, IN, 1980-84. I was a teaching assistant in Calculus, which involved conducting recitation sessions, writing and grading of homework assignments and exams, and I lectured in college algebra, trigonometry and analytic geometry, wrote exams, assigned homework, and supervised undergraduate paper graders. In the Dept. of Foreign Languages I lectured in Russian language and culture; wrote exams, assigned and graded homework.

**Molecular Biologist**, Dept. of Experimental Pathology, University of Kentucky Medical Center, 1979-80. I

provided technical assistance in the study of an avian cancer virus: molecular biological methods, avian cell tissue culture.

**Publications**

*Inventions*

**Zhou, CE**, A Zemla, D Roe, J Schoeinger. Countermeasures to Common Motifs (CCM): a computational infrastructure for discovery and characterization of druggable targets on proteins and for down-selection of drug candidates (LLNL case no. IL-11683; U.S. patent application filed 12 June 2007)

**Zhou, CE**, A Zemla. Structure-based analysis for identification of protein signatures: cuScore (LLNL case no. IL-11677; U.S. patent application filed 16 April 2007; 11/735,972).

**Zhou, CE**, A Zemla, M Lam. Sequence-based analysis for identification of protein signatures: pScore (LLNL case no. IL-11678; U.S. patent application filed 16 April 2007; 11/735,981).

Zemla, A, **C Zhou**, J Smith, and M Lam. STRALCP – Structure alignment-based clustering of proteins. (LLNL case no. IL-11696; provisional patent filed 2006)

**Zhou, CE**, A Zemla, M Lam, J Smith, E Vitalis, S Gardner, T Kuczmarski, T Slezak, C Torres, D Roe, J Schoeniger, M Young. Protein Signature Evaluation (PSE) system: a high-throughput, whole-proteome, bio-informatics approach to protein target and signatures discovery and characterization (LLNL case no. IL-11705; provisional patent filed 2006).

Zemla, AT, **CE Zhou**, M Lam, J Smith, E Pardes. cSpan structure-sequence-based analysis for identification of conservation regions in proteins (LLNL case no. IL-11776; record of invention 2007).

Roter, A, H Kataoka, R Troetschler, S Maeda, F Enderlin, **C Ecale**, S Kramer and D Schooley. An insect virus insecticide with enhanced efficacy containing a gene encoding eclosion hormone of *Manduca sexta* (Sandoz Crop Protection, disclosure of invention, 1987).

Roter, A, H Kataoka, R Troetschler, S Maeda, F Enderlin, **C Ecale**, S Kramer and D Schooley. An insect virus insecticide with enhanced efficacy containing a gene encoding diuretic hormone of *Manduca sexta* (Sandoz Crop Protection, disclosure of invention, 1987).

*Refereed Journal Articles*

Zemla, A, B Geisbrecht, J Smith, M Lam, B Kirkpatrick, M Wagner, T Slezak, and **CE Zhou**. STRALCP: structure alignment-based clustering of proteins. *Nucleic Acids Research* (in press).

Zemla, A and **CE Zhou**. Structural re-alignment in an immunogenic surface region of ricin A chain. *Bioinformatics and Biology Insights* (in press).

**Zhou, CE**, J Smith, M Lam, A Zemla, M Dyer, and T Slezak. 2006. MvirDB—A microbial database of protein toxins, virulence factors, and antibiotic resistance genes for bio-defense applications. *Nucleic Acids Research Database Issue 2007* doi:10.1093/nar/gk1791; NAR Online Molecular Biology Database Collection No. 996.

**Zhou, CE**, M Lam, J Smith, A Zemla, M Dyer, T Kuczmarski, E Vitalis, and T Slezak. 2006. MannDB: a microbial database of automated protein sequence analyses and evidence integration for protein characterization. *BMC Bioinformatics* 7:459.

**Zhou, CE**, A Zemla, D Roe, M Young, M Lam, JS Schoeniger, and R Balhorn. 2005. Computational approaches for identification of conserved/unique binding pockets in the A chain of ricin. *Bioinformatics* 21:3085-3096.

Zemla, A, **C Ecale Zhou**, T Slezak, T Kuczmarski, D Rama, C Torres, D Sawicka, and D Barsky. 2005. AS2TS system for protein structure modeling and analysis. *Nucleic Acids Research* 33: W111-W115.

Gardner, SN, Kuczmarski, TA, **Zhou, CE**, Lam, MW, Slezak, TR. 2005. A system to assess genome sequencing needs for viral protein diagnostics and therapeutics. *Journal of Clinical Microbiology*. 43:1807-1817.

Taniai, K, **CE Zhou**, DG Lee, S Maeda, and BD Hammock. 2005. Expression of a non-secreted form of juvenile hormone esterase in a baculovirus. *Japan Agricultural Research Quarterly* 39:11-18.

Gocns, SD, S Botero, A Zemla, **C Ecale Zhou**, and ML Perdue. 2004. Bovine enterovirus type 2. Complete genomic sequence and molecular modeling of the reference strain and a wild type isolate from endemically infected US cattle. *Journal of General Virology* 85:3195-3203.

Slezak T, T Kuczmarski, L Ott, C Torres, D Mederos, J Smith, B Truitt, N Mulakken, M Lam, E Vitalis, A Zemla, **C Zhou**, S Gardner. 2003. Comparative genomics tools applied to bioterrorism defense. *Briefings in Bioinformatics*, Vol. 4, No. 2, 133-149.

**Ecale Zhou, C**, E Ammar, H Sheta, S Kelley, M Polek, and D Ullman. 2002. Citrus tristeza virus ultrastructure and associated cytopathology in *Citrus sinensis* and *Citrus aurantifolia*. *Canadian Journal of Botany* 80: 512-525.

Dehal, P, P Predki, A Olsen, A Kobayashi, P Folta, S Lucas, M Land, A Terry, **C Zhou**, S Rash, Q Zhang, L Gordon, J Kim, C Elkin, M Pollard, P Richardson, D Rokhsar, E Uberbacher, T Hawkins, E Branscomb, and L

Stubbs. 2001. Human chromosome 19 and related regions in mouse: conservative and lineage-specific evolution. *Science* 293: 104-111.
Katsuma S, D Deng, **C Zhou**, M Iwanaga, Y Noguchi, M Kobayashi, and S Maeda. 2000. Identification of novel residues involved in nuclear localization of a baculovirus polyhedrin protein. *Virus Genes* 21: 233-240.
Katsuma, S, Y Noguchi, **C Ecale Zhou**, M Kobayashi, and S Maeda. 1999. Characterization of the 25K FP gene of the baculovirus BmNPV: implications for postmortem host degradation. *J. General Virology* 80: 783–791.
**Ecale Zhou, C**, R Ko and S Maeda. 1998. Polyhedron-like inclusion body formation by a mutant nucleopolyhedrovirus expressing the granulin gene from a granulovirus. *Virology* 240: 282–294 (cover).
**Ecale, C** and E Backus. Phloem injury and repair following potato leafhopper feeding on alfalfa stems. 1998. *Candian Journal of Botany* 77: 537-547.
Gomi, S, **C Ecale Zhou** and S Maeda. 1997. Characterization of late expression factor gene homologues of the baculovirus, BmNPV. *Virology* 230: 35–47.
**Ecale, C** and E Backus. 1995. Time course of anatomical changes to vascular tissues of alfalfa, *Medicago sativa* L., from probing injury by the potato leafhopper, *Empoasca fabae* (Harris). *Canadian Journal of Botany* 73: 288–298.
**Ecale, C** and E Backus. 1995. Mechanical and salivary aspects of potato leafhopper probing damage to alfalfa stems. *Entomologia Experimentalis et Applicata* 77: 121–132.
**Ecale, C** and E Backus. 1994. A characteristic body posture of the potato leafhopper is correlated with probing. *Journal of Entomological Science* 23(4): 490–495.
**Ecale, C.** 1990. "The Happiest Day," translation from the original Russian of "Samyj schastlivy den'," by Viktoria Tokareva. *The Massachusetts Review* 31(3): 327–37.

***Proceedings Papers***
**Ecale, C** and E Backus. 1993. Use of videomicrography to standardize potato leafhopper probing in the study of hopperburn, pp. 72-73. In: *Proceedings of the 8th International Auchenorrhyncha Congress*, Delphi, Greece.
**Ecale, C** and E Backus. 1993. Salivary secretions of the potato leafhopper into agarose artificial diet and stems of alfalfa. In: *Proceedings of the X Annual Research and Creative Activities Forum*, University of Missouri.
**Ecale, C** and E Backus. 1992. Mechanical and salivary effects of potato leafhopper feeding in relation to hopperburn in alfalfa, p. 222. In: *Proceedings of the XIX International Congress of Entomology*, Beijing, China.

***Papers Submitted or in Preparation***
**Zhou, CE**, A Zemla, J Smith, M Lam, R Balhorn, and T Slezak. Identification of protein signatures for West Nile virus envelope glycoprotein using the Protein Signatures Evaluation system. (in preparation)

---

**Invited Talk**

Zhou, CE, "MvirDB—A Virulence Data Warehouse for Bio-defense Applications", DIA DTT BW S&T Symposium, June 19-21, 2007, MIT Lincoln Lab, Lexington, MA.

---

**Education**

**B.S., Computer Science**, California State Univ., Sacramento, CA, 1999.
**Ph.D., Biological Science/Entomology**, University of Missouri, Columbia, MO, 1993.
**B.S., Chemistry**, Purdue University, W. Lafayette, IN, 1979.

*Extras*
UCLA Technical Management Program
LLNL Emerging Leaders Program
Oracle 8i Database Administration series
Copyediting, Univ. California-Berkeley Extension

---

DOE Security Clearance Level Q